

# Guaranteeing Spoof-Resilient Multi-Robot Networks

Swarun Kumar<sup>†</sup> Stephanie Gil<sup>†</sup> Mark Mazumder<sup>\*</sup> Dina Katabi<sup>†</sup> Daniela Rus<sup>†</sup>

<sup>†</sup>Massachusetts Institute of Technology

{swarun, sgil, dk, rus}@mit.edu

<sup>\*</sup>MIT Lincoln Laboratory

mazumder@ll.mit.edu

**Abstract**—Multi-robot networks use wireless communication to provide wide-ranging services such as aerial surveillance and unmanned delivery. However, effective coordination between multiple robots requires trust, making them particularly vulnerable to cyber-attacks. Specifically, such networks can be gravely disrupted by the Sybil attack, where even a single malicious robot can spoof a large number of fake clients. This paper proposes a new solution to defend against the Sybil attack, without requiring expensive cryptographic key-distribution. Our core contribution is a novel algorithm implemented on commercial Wi-Fi radios that can “sense” spoofers using the physics of wireless signals. We derive theoretical guarantees on how this algorithm bounds the impact of the Sybil Attack on a broad class of robotic coverage problems. We experimentally validate our claims using a team of AscTec quadrotor servers and iRobot Create ground clients, and demonstrate spoiler detection rates over 96%.

## I. INTRODUCTION

Multi-robot networks rely on wireless communication to enable a wide range of tasks and applications: coverage [26, 5, 29], disaster management [6], surveillance [3], and consensus [25] to name a few. The future promises an increasing trend in this direction, such as delivery drones which transport goods (e.g. Amazon Prime Air [1]) or traffic rerouting algorithms (e.g. Google Maps Navigation) that rely on broadcasted user locations to achieve their goals. Effective coordination, however, requires trust. In order for these multi-robot systems to perform their tasks optimally, transmitted data is often assumed to be accurate and trustworthy; an assumption that is easy to break. A particularly challenging attack on this assumption is the so-called “Sybil attack.”

In a Sybil attack a malicious agent can generate (or spoof) a large number of false identities to gain a disproportionate influence in the network.<sup>1</sup> These attacks are notoriously easy to implement [31] and can be detrimental for multi-robot networks. An example of this is coverage, where an adversarial client can spoof a cluster of clients in its vicinity in order to create a high local demand, in turn denying service to legitimate clients (see Figure 1). Although there is a vast body of literature dedicated to cybersecurity in general multi-node networks (e.g. a wired LAN), the same is not true for multi-robot networks [14, 28], leaving them largely vulnerable to these types of attacks. This is because many characteristics unique to robotic networks make security more challenging; for example, traditional key passing or cryptographic authentication is difficult to maintain due to the highly dynamic and distributed nature of multi-robot teams where clients often enter and exit the network.

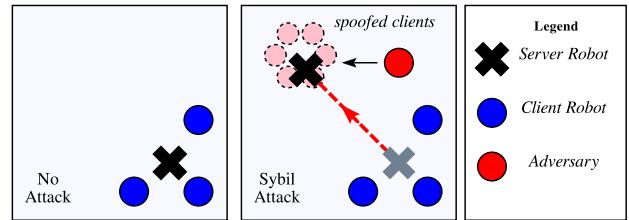


Fig. 1: **Sybil Attack on Coverage:** A server robot provides locational coverage to legitimate clients when no attack is present. In a Sybil attack, an adversary spoofs fake clients to draw away coverage from the legitimate clients.

This paper addresses the challenge of guarding against Sybil attacks in multi-robot networks. We focus on the general class of problems where a group of server robots coordinate to provide some service using the broadcasted locations of a group of client robots. Our core contribution is a novel algorithm that analyzes the received wireless signals to detect the presence of spoofed clients spawned by adversaries. We call this a “virtual spoiler sensor” as we do not use specialized hardware nor encrypted key exchange, but rather a commercial Wi-Fi card and software to implement our solution. Our virtual sensor leverages the rich physical information already present in wireless signals. At a high level, as wireless signals propagate, they interact with the environment via scattering and absorption from objects along the traversed paths. Carefully processed, these signals can provide a unique signature or “spatial fingerprint” for each client, measuring the power of the signal received along each spatial direction (Fig. 2). Unlike message contents such as reported IDs or locations which adversaries can manipulate, spatial fingerprints rely on physical signal interactions that cannot be exactly predicted [12, 22].

Using these derived fingerprints, we show that a confidence weight,  $\alpha \in (0, 1)$  can be obtained for each client in the network. We prove that these confidence weights have a desirable property where legitimate clients have an expected confidence weight close to one, while spoofed clients will have an expected confidence weight close to zero. A particularly attractive feature of confidence weight  $\alpha$  is that it can be readily integrated as a per-client weighting function into a wide variety of multi-robot controllers. More importantly, the analytical bounds on these weights can provably limit the ill-effects of spoofers on the performance of these controllers. This paper demonstrates this capability in the context of the well-known locational coverage algorithm [5, 29].

We provide an extensive experimental evaluation of our theoretical claims using a heterogeneous team of air/ground

<sup>1</sup>Please refer to [7, 24] for a detailed treatment of this class of cyber attacks.

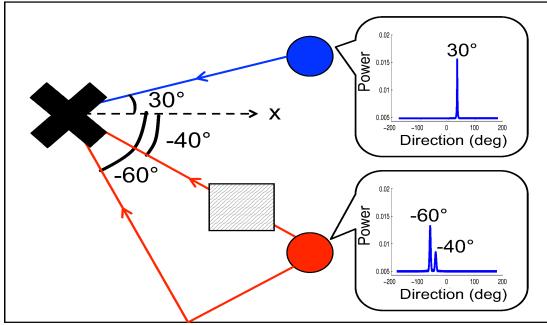


Fig. 2: **Spatial Fingerprints:** A quadrotor server measures the directional signal strength of each client (here, simplified to 2-D). The blue client has one line-of-sight peak; the other, 2 signal paths.

robots consisting of two AscTec Hummingbird platforms and ten iRobot Create platforms. We conduct our experiments in general indoor settings with randomly placed clients and demonstrate a spoofer detection rate of 96%. For the case of coverage we find that the converged positions of the service robots is on average 3 cm from optimal even when more than 75% of total clients in the network are spoofed.

*Contributions of this paper:* We develop a virtual sensor for spoofing detection which provides performance guarantees in the presence of Sybil attacks and is applicable to a broad class of problems in distributed robotics. We show that the influence of spoofers is analytically bounded under our system in a coverage context, where each robotic node providing coverage remains within a radius of its position in the absence of an attack. Our theoretical results are validated extensively through experiments in diverse settings.

## II. RELATED WORK

The problem of Sybil attacks has been studied in general multi-node, often static, networks, and many tools have been developed for these settings. Past work falls under three categories: (1) Cryptographic Authentication Schemes can be used to prevent Sybil Attacks (See Table 7 in [37]). These require trusted central authorities and computationally expensive distributed key management, to account for dynamic clients that enter and leave the network [37]. (2) Non-cryptographic techniques in the wireless networking community leverage wireless physical-layer information to detect spoofed client identities or falsified locations [15, 40, 38, 39]. These rely on bulky and expensive hardware like large multi-antenna arrays, that cannot be mounted on small robotic platforms. (3) Recent techniques have attempted to use wireless signal information like received signal strength (RSSI) [35, 27] and channel state information [21]. Such techniques need clients to remain static, since mobility can cause wireless channels to fluctuate rapidly [2]. In addition, they are susceptible to power-scaling attacks, where clients scale power differently to imitate different users. In sum, the above systems share one or more of the following characteristics making them ill-suited to multi-robot networks: (1) Require computationally-intensive key management; (2) Rely on bulky and expensive hardware; (3) Assume static networks. Indeed past work has highlighted

the gravity and apparent sparsity of solutions to cyber-security threats in multi-robot networks [14, 28, 4].

Unlike past work, our solution has three attributes that particularly suit multi-robot networks: (1) It captures physical properties of wireless signals and therefore does not require distributed key management [37]. (2) It relies on cheap commodity Wi-Fi radios, unlike hardware-based solutions [38, 40]. (3) It is robust to client mobility and power-scaling attacks.

Finally, our system builds on Synthetic Aperture Radar (SAR) to construct signal fingerprints [8]. SAR has been widely used for radar imaging [8, 16] and indoor positioning [18, 17, 34, 11]. In contrast, this paper builds upon SAR to provide cyber-security to multi-robot networks. In doing so, it provides theoretical security guarantees that are validated experimentally. These integrate readily with performance guarantees of existing multi-robot controllers, like the well-known robotic coverage controllers [5, 29] as shown in Sec. §VI.

## III. PROBLEM STATEMENT

This paper focuses on problems where the knowledge of agent positions facilitates some collaborative task. Specifically, it assumes two groups of agents, “clients” requiring some type of location-based service such as coverage or goods delivery and “servers” whose positions are optimized in order to provide the service to its clients. Let  $P := \{p_1, \dots, p_c\}$  denote the client positions in  $\mathbb{R}^3$ . Let  $X := \{x_1, \dots, x_m\}$  be the positions of the servers in  $\mathbb{R}^3$  and the notation  $[m] = \{1, \dots, m\}$  denote their indices. We consider the case where a subset of the clients,  $S \subset P, |S| = s$  are “spoofed” clients.

*Definition 3.1 (Spoofed Client):* A spoofed client is a client with a reported position  $p \in \mathbb{R}^3$ , different from its ground truth position  $\bar{p} \in \mathbb{R}^3$  beyond a given error tolerance  $\sigma > 0^\circ$ , with respect to any server robot in the network. Specifically, let  $e_l = \angle((x_l - p), (x_l - \bar{p}))$  be the angle between a server at position  $x_l$  and a client at reported position  $p$ . Then a client whose angle  $e_l$  exceeds  $\sigma$  degrees, to any server  $x_l \in X$ , is considered spoofed. Clients who are not spoofed are “legitimate” clients.<sup>2</sup> A single adversarial client can generate an arbitrary number of spoofed clients, each with fabricated positions.

**Threat Model:** Our threat model considers one or more adversarial robot clients with one Wi-Fi antenna each. The adversaries can be mobile and scale power on a per-packet basis. We only consider adversarial clients.<sup>3</sup> Adversarial clients perform the “Sybil Attack” to forge packets emulating  $s$  non-existent clients, where  $s$  can exceed the number of legitimate clients. More formally:

*Definition 3.2 (Sybil Attack):* Define a network of clients and servers as  $P \cup X$ , where a subset  $S$  of the clients are spoofers, such that  $P = S \cup \tilde{S}$ . We assume that set  $P$  is known but knowledge of which clients are spoofed (i.e., in  $S$ ) is unknown. This attack is called a “Sybil Attack.”

To counter the Sybil attack, this paper has two objectives. First, we find a relation capturing directional signal strength

<sup>2</sup>Sec. §VII examines spoofed clients co-aligned with legitimate clients.

<sup>3</sup>The case of adversarial server robots is left for future work although many of the concepts in the current paper are extensible to this case as well.

between a client  $i$  and a server  $l$ . We seek a mapping  $F_{il} : [0, \frac{\pi}{2}] \times [0, 2\pi] \mapsto \mathbb{R}$  such that for any 3D direction  $(\theta, \phi)$  defined in Fig. 4, the value  $F_{il}(\theta, \phi)$  is the power of the received signal from client  $i$  along that direction. Using this mapping, or “fingerprint”, our first problem is to derive a *confidence weight* whose expectation is provably bounded near 1 for legitimate clients and near 0 for spoofed clients. Further, we wish to find these bounds analytically from problem parameters like the signal-to-noise ratio of the received wireless signal. We summarize this objective as Problem 1 below:

**Problem 1: Spoof Detection** Let  $\mathcal{F}_i$  be the set of fingerprints measured from all clients  $j \in [c]$  and servers  $l \in [m]$  in the neighborhood,  $\mathcal{N}_i$ , of client  $i$ .<sup>4</sup> Here, a neighborhood of client  $i$ ,  $\mathcal{N}_i$ , are all agents that can receive Wi-Fi transmissions sent by client  $i$ . Using  $\mathcal{F}_i$ , derive a confidence weight  $\alpha_i(\mathcal{F}_i) \in (0, 1)$  and a threshold  $\omega_i(\sigma_i^2) > 0$  where  $\sigma_i^2$  represents error variances such as the signal-to-noise ratio that are assumed to be given. Find  $\omega_i(\cdot)$  to have the provable property of differentiating spoof clients whereby spoof clients are bounded below this threshold, i.e.  $E[\alpha_i] \leq \omega$ , and legitimate clients are bounded above this threshold  $E[\alpha_i] \geq 1 - \omega$ .

Our second objective is to apply our spoof detection method to multi-robot control problems. We consider the well-known coverage problem in [5, 29]. We show that by integrating the confidence weight from Problem 1, we can analytically bound the error in performance caused by spoofed clients in the network. We consider the coverage problem where an importance function is defined over an environment and where the positions of the clients correspond to peaks in the importance function. Here, servers position themselves to maximize their proximity to these peaks, to improve their coverage over client robots. If  $C_V = \{x_1^*, \dots, x_m^*\}$  is the set of server positions optimized by the coverage controller with zero spoofers, we wish to guarantee that server positions optimized with spoofers present,  $C_{V_\alpha}$ , is “close” to  $C_V$ . We state this second objective more specifically as Problem 2 below:

### Problem 2: Sybil-resillience in Multi-Robot Coverage

Consider a locational coverage problem where an importance function  $\rho(q) > 0$  is defined over an environment  $\mathcal{Q} \subset \mathbb{R}^3$  and  $q \in \mathcal{Q}$ . Specifically, consider an importance function that can be decomposed into terms,  $\rho_i(q)$ , depending on each client’s position,  $i \in [c]$  (for example, each client position corresponds to a peak), i.e.  $\rho(q) = \rho_1(q) + \dots + \rho_c(q)$ . Let  $C_V = \{x_1^*, \dots, x_m^*\}$  be the set of server positions returned by an optimization of  $\rho(q)$  over  $X$ , where there are zero spoofed clients in the network. Under a Sybil attack, let  $C_{V_\alpha} = \{x_1, \dots, x_m\}$  be the set of server positions returned by an optimization of an  $\alpha$ -modified importance function  $\rho(q) = \alpha_1\rho_1(q) + \dots + \alpha_c\rho_c(q)$  where the importance weight terms  $\alpha_i$  satisfy the bounds stated in Problem 1. We wish to

<sup>4</sup>The more servers there are to sense the wireless transmissions from client  $i$  (i.e. larger neighborhoods  $\mathcal{N}_i$ ), the easier it becomes to detect whether client  $i$  is being spoofed. But we note that even with a single server this determination can be made. A theoretical treatment of this point can be found in Sec. §V and experimental results in Sec. §VII-A use as little as one server in the system.

find an  $\epsilon(\mathcal{P}) > 0$  such that the set  $C_{V_\alpha}$  is within a distance  $\epsilon(\mathcal{P})$  to  $C_V$ .  $C_{V_\alpha}$  is within a distance  $\epsilon(\mathcal{P})$  to  $C_V$  if  $\forall x \in C_{V_\alpha}$  there exists a unique  $y \in C_V$  where  $\text{dist}(x, y) < \epsilon(\mathcal{P})$ . Here,  $\mathcal{P}$  is a set of problem parameters that we wish to find.

Intuitively, solutions to Problem 2 guarantee that under a Sybil attack, all server positions computed using an  $\alpha$ -modified coverage controller are within a computable distance  $\epsilon(\mathcal{P})$  from their optimal positions (i.e. in the absence of spoofers). Sec. §VI derives a closed-form for  $\epsilon(\mathcal{P})$  and shows the set  $\mathcal{P}$  of problem parameters to be the number of spoofers, the footprint of the environment covered, and signal noise.

### IV. FINGERPRINTS TO DETECT MALICIOUS CLIENTS

In this section, we develop unique client fingerprints based on the physics of their wireless signals. Specifically, we leverage wireless channels  $h$ , complex numbers measurable on any wireless device characterizing the attenuation in power and phase rotation signals experience as they propagate over the air. These channels also capture the fact that wireless signals are scattered by the environment, arriving at the receiver over (potentially) several different paths [33]. Fig. 3 is an example 2D schematic of a wireless signal traversing from a client robot to a server robot arriving along two separate paths: one attenuated direct path at  $40^\circ$  and one reflected at  $60^\circ$ . If the server robot had a directional antenna, it could obtain a full 3D profile of power of the received signal (i.e.  $|h|^2$ ) along every spatial direction. This would be an ideal “spatial fingerprint” since such a profile is 1) highly position dependent and 2) not controllable by the sender (since the occurrence of individual paths is due to reflectors in the environment).

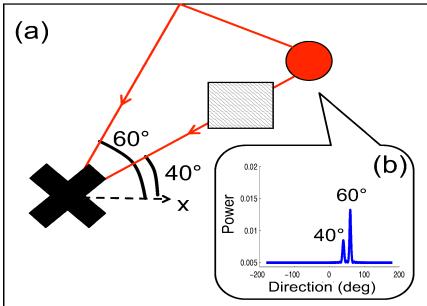
Unfortunately directional antennas are composed of large arrays of many antennas that are too bulky for small agile robot platforms. Luckily, a well-known technique called Synthetic Aperture Radar [8] (SAR) can be used to emulate such an antenna using a commodity Wi-Fi radio. Its key idea is to use small local robotic motion, such as spinning in-place, to obtain multiple snapshots of the wireless channel that are then processed like a directional array of antennas. SAR can be implemented using a well-studied signal processing algorithm called MUSIC [13] to obtain spatial fingerprints at each server robot.

Mathematically, we obtain a spatial fingerprint for each wireless link between a server  $l$  and client  $i$  as a matrix  $F_{il} : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ . For each spatial path represented as  $(\theta, \phi)$  (see Fig. 4),  $F_{ij}$  maps to a scalar value representing the signal power received along that path. More formally:

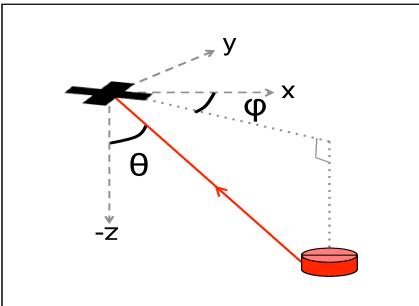
$$F_{il}(\phi, \theta) = 1 / |Eig_n(\hat{h}_{il}\hat{h}_{il}^\dagger)e^{\sqrt{-1}\Psi_{il}(\phi, \theta)}|^2 \quad (1)$$

Where  $\hat{h}_{il}$  is a vector of the ratio of wireless channel snapshots between two antennas mounted on the body of the server  $l$  and  $\Psi_{il}(\phi, \theta) = \frac{2\pi r}{\lambda} \cos(\phi - \mathbf{B}_1) \sin(\theta - \Gamma_1)$ ,  $\lambda$  is the wavelength of the signal and  $r$  is the distance between the antennas,  $Eig_n(\cdot)$  are noise eigenvectors,  $(\cdot)^\dagger$  is conjugate transpose, and  $k$  is the number of signal eigenvectors, equal to the number of paths.

While our above formulation is derived from MUSIC [13], it varies in one important way: While MUSIC uses a single-antenna channel snapshot  $h_{il}$ , we use the channel ratio  $\hat{h}_{il} =$



**Fig. 3: Example Signal Fingerprint:** (a) A server (black) receives signal from client (red) on 2 paths: direct along  $40^\circ$  attenuated by obstacle (shaded) and reflected by wall along  $60^\circ$ . (b) shows corresponding fingerprint with peaks at  $40^\circ$  and  $60^\circ$  with heights corresponding to their relative attenuations.



**Fig. 4: 3-D Angles:** The figure depicts the notation for the azimuthal angle  $\phi$  and polar angle  $\theta$  for the direct path from a ground client (red) to aerial server robot (black) in 3-dimensions. More generally, the set of all angles between client  $i$  and server  $l$  are denoted as  $\Phi_{il}, \Theta_{il}$  respectively.

Symbol	Meaning
$m, c, s$	No. of servers, clients, spoofers
$p_i, x_l$	Position of client $i$ / server $l$
$F_{il}, k$	Fingerprint of $i$ at $l$ , $k$ peaks
$\hat{h}_{il}$	$M \times 1$ channel ratios of $i$ to $l$
$f(\cdot; \mu, \sigma^2)$	PDF of normal distribution
$g(\cdot; \mu, \sigma^2)$	$\min(1, \sqrt{2\pi}f(x; \mu, \sigma^2))$
$\kappa$	Constant $= ((\sqrt{2} + \sqrt{\pi})/\pi)^2$
$\alpha_i, \beta_i$	confidence, honesty metric of $i$
$\gamma_{ij}$	Similarity metric of client $i, j$
SNR	Signal-to-noise ratio
RSSI	Received Signal Strength
$\sigma_\theta^2, \sigma_\phi^2$	Variance in peak shifts of $F_{il}$
$\hat{\sigma}_\theta^2, \hat{\sigma}_\phi^2$	$\sigma_\theta^2, \sigma_\phi^2$ plus measurement error
$C_{VL}, C_{V_\alpha}$	Coverage centroid of optimal, our system; error $\vec{e}$ within $\epsilon$
$L(Q), \rho(q)$	Footprint, Mass function

**Fig. 5: Table of Most Common Notations**

$h_{1_{il}}/h_{2_{il}}$  between two antennas. This modification provides resilience to intentional power scaling by the sender since scaling his transmit power by  $\chi$  yields a measured ratio  $\hat{h}_{il} = \chi h_{1_{il}}/(\chi h_{2_{il}})$ ; a value unaffected by power scaling.

## V. CONSTRUCTING A CLIENT CONFIDENCE WEIGHT

In this section, we leverage the unique client fingerprints  $F_{il}(\phi, \theta)$  for each user  $i$  relative to the robotic server  $l$  to generate a confidence weight  $\alpha_i \in [0, 1]$  on whether client  $i$  is legitimate or not.  $\alpha_i$  approaches 1 if client  $i$  is suspected to be legitimate, and 0 otherwise. It is defined as the product of two components: the honesty metric  $\beta_i$ , and the similarity metric  $\gamma_{ij}$ . These components measure the likelihood of the following events: (1) The honesty metric  $\beta_i$  captures whether client  $i$  is honest about the position it reports. Specifically,  $\beta_i$  approaches 1 if the client  $i$ 's reported location has a corresponding peak in its fingerprint; (2) The similarity metric  $\gamma_{ij}$  captures whether clients  $i$  is identical to another client  $j$  (i.e. is a spoofed client).  $\gamma_{ij}$  approaches 1 if client  $i$ 's fingerprint appears identical to another client  $j$ . Mathematically:

$$\begin{aligned} \alpha_i &= \beta_i \prod_{j \neq i} (1 - \gamma_{ij}) \quad \text{where, } \beta_i = \prod_l \mathcal{L}(i \text{ is at } (\phi_{il}, \theta_{il}) | F_{il}) \\ \gamma_{ij} &= \prod_l \mathcal{L}(i \text{ spoofs } j | F_{il}, F_{jl}) \end{aligned} \quad (2)$$

Here,  $\mathcal{L}(\cdot)$  denotes likelihood of an event and  $(\phi_{il}, \theta_{il})$  are the expected direction of client  $i$ , from its reported location.

**Defining Honesty and Similarity Metric:** To define the honesty metric  $\beta_i$  and similarity metric  $\gamma_{ij}$  precisely, one must account for the effect of noise. Specifically, both these metrics inspect the locations of peaks of client fingerprints. In practice however, these peaks may have slight shifts owing to noise. This means that any comparison between peak locations must permit some variance due to these shifts. Fortunately, noise in wireless environments can be modeled closely as additive white-Gaussian [33]. As the following lemma shows, this results in shifts in peaks that are also Gaussian, meaning that their variance is easy to model and account for. More formally, the lemma states that the shifts are normally distributed with

zero mean and well-defined variance, based on the signal-to-noise ratio (SNR) of the wireless medium:

*Lemma 5.1:* Let  $\Delta\theta_i, \Delta\phi_i$  denote the error between the azimuthal and polar angle of the uncorrelated  $i^{\text{th}}$  path of a (potentially multipath) source and the corresponding angles of the (local) maximum in the profile  $F(\phi, \theta)$ , gathered over a large number of uniformly gathered packets (i.e. SAR snapshots) for  $\theta \in (10^\circ, 80^\circ)$ . Then  $\Delta\theta_i$  and  $\Delta\phi_i$  are normally distributed with a mean 0, and expected variance  $\sigma_\phi^2$  and  $\sigma_\theta^2$ :

$$\sigma_\theta^2 = \sigma_\phi^2 = 9\lambda^2/(8M\pi^2 r^2 SNR)$$

Where,  $\lambda$  is the wavelength of the signal,  $SNR$  is the signal-to-noise ratio in the network<sup>5</sup>,  $M$  is the number of packets per-rotation, and  $r$  is the distance between the antennas.  $\square$

The above lemma follows from well-known Cramer-Rao bounds [23, 10, 9] shown previously for linear antenna movements in SAR [32] but readily extensible to circular rotations (proof in supplementary material). Using this lemma, we can define the honesty metric  $\beta_i$  as the likelihood that the client is at its reported location, subject to this Gaussian error and additional measurement error in reported locations.

*Definition 5.2:* ( $\beta_i$ ) Let  $\phi_{F_{il}}$  and  $\theta_{F_{il}}$  denote the closest local maximum in  $F_{il}(\phi, \theta)$  to  $(\phi_{il}, \theta_{il})$ . We denote  $\hat{\sigma}_\phi^2$  and  $\hat{\sigma}_\theta^2$  as the variances in angles  $\sigma_\phi^2$  and  $\sigma_\theta^2$  plus any variance due to measurement error of reported locations that can be calibrated from device hardware. We define  $\beta_i$  for client  $i$  as:

$$\beta_i = \prod_l g(\phi_{il} - \phi_{F_{il}}; 0, \hat{\sigma}_\phi^2) \times g(\theta_{il} - \theta_{F_{il}}; 0, \hat{\sigma}_\theta^2) \quad (3)$$

Where  $g(x; \mu, \sigma^2) = \min(1, \sqrt{2\pi}f(x; \mu, \sigma^2))$  is a normalized Gaussian PDF  $f(x; \mu, \sigma^2)$  with mean  $\mu$  and variance  $\sigma^2$ .  $\square$

Similarly, the similarity metric  $\gamma_{ij}$  is the likelihood that two clients share identical peaks in their fingerprints, subject to Gaussian shift in their respective peaks from Lemma. 5.1.

*Definition 5.3:* ( $\gamma_{ij}$ ) Let  $(\Phi_{il}, \Theta_{il})$  and  $(\Phi_{jl}, \Theta_{jl})$  denote the ordered set of local maxima in profiles  $F_{il}$  and  $F_{jl}$ . We

<sup>5</sup>For clarity, we drop dependence on  $i, l$  for SNR,  $\sigma_\theta$  and  $\sigma_\phi$

define  $\gamma_{ij}$  for client  $i$  relative to client  $j$  as:

$$\gamma_{ij} = \prod_{\phi_i \in \Phi_{il}, \phi_j \in \Phi_{jl}} g(\phi_i - \phi_j; 0, 2\sigma_\phi^2) \prod_{\theta_i \in \Theta_{il}, \theta_j \in \Theta_{jl}} g(\theta_i - \theta_j; 0, 2\sigma_\theta^2) \quad (4)$$

Where  $g(\cdot; \mu, \sigma^2)$  is as defined in Definition 5.2.  $\square$

**Defining the Confidence Weight:** We notice that Eqn. 2, 3 and 4 fully define  $\alpha_i$  for each client  $i$ . In summary, the confidence weight is computed in three steps: (1) Obtain the client fingerprint using SAR on wireless signal snapshots. (2) Measure the variance of peak locations of these client fingerprints using their Signal-to-Noise Ratio. (3) Compute the similarity and honesty metrics using their above definitions to obtain the confidence weight. Algorithm 1 below summarizes the steps to construct  $\alpha_i$  for a given client  $i$ .

---

#### Algorithm 1 Algorithm to Compute Client Confidence Weight

---

```

▷ Input: Ratio of Channels  $\hat{h}_{il}$  and SNR
▷ Output: Confidence Weight,  $\alpha_i$  for client  $i$ 
▷ Step (1): Measure fingerprints for client  $i$ 
for  $l = 1, \dots, m$  do
    for  $\phi \in \{0^\circ, \dots, 360^\circ\}; \theta \in \{0^\circ, \dots, 360^\circ\}$  do
        Find  $F_{il}(\phi, \theta)$  using a single spin to get  $\hat{h}_{il}$  (Eqn. 1)
    end for
end for
▷ Step (2): Measure variances in peak locations using SNR
 $\sigma_\theta^2 = \sigma_\phi^2 =$  Apply Lemma 5.1 SNR
▷ Step (3): Find honesty, similarity and confidence weight
 $\beta_i =$  Apply Defn. 5.2 using  $\sigma_\theta^2, \sigma_\phi^2$ , peaks of  $F_{il}$ 
for  $j = \{1, \dots, c\} \setminus \{i\}$  do
     $\gamma_{ij} =$  Apply Defn. 5.3 using  $\sigma_\theta^2, \sigma_\phi^2$ , peaks of  $F_{il}, F_{jl}$ 
end for
 $\alpha_i = \beta_i \prod_{j \neq i} (1 - \gamma_{ij})$ 

```

---

We now present our main result that solves Problem 1 in the problem statement (Sec. §III). The following theorem says the expected  $\alpha_i$ 's of legitimate nodes approach 1, while those of spoofers approach 0, allowing us to discern them under well-defined assumptions: (A.1) The signal paths are independent. (A.2) Errors in azimuth and polar angles are independent. (A.3) The clients transmit a large number of packets.

**Theorem 5.4:** Consider a network with  $m$  servers and  $c$  clients. A new client  $i$  either: 1) spoofs  $s$  clients reporting a random location, potentially scaling power, or; 2) is a uniformly randomly located legitimate client. Let  $\alpha_{spoof}, \alpha_{legit}$  be the confidence weights in either case. Assume that the client obtains its signals from servers along  $k$  paths. Under A.1-A.3, the expected  $\alpha_{spoof}, \alpha_{legit}$  are bounded by:

$$\begin{aligned} E[\alpha_{spoof}] &\leq \left[ \sqrt{\hat{\sigma}_\theta \hat{\sigma}_\phi} \kappa \right]^m [2mk\sigma_\theta\sigma_\phi]^s \\ E[\alpha_{legit}] &\geq 1 - cm\hat{\sigma}_\theta\hat{\sigma}_\phi \left[ \sqrt{2\sigma_\theta\sigma_\phi} \kappa \right]^{mk} \end{aligned} \quad (5)$$

Where  $\kappa = ((\sqrt{2} + \sqrt{\pi})/\pi)^2$ ,  $\sigma_\theta, \sigma_\phi, \hat{\sigma}_\theta, \hat{\sigma}_\phi$  are the variances defined in Lemma 5.1 that depend on signal-to-noise ratio (the latter include measurement error in reported locations).

**Proof Sketch:** To give some intuition on why the theorem holds, we provide a brief proof sketch (detailed proof is available in the supplementary material). To begin with, notice from their definitions that both the honesty metric  $\beta_i$  and confidence metric  $\gamma_{ij}$  inspect peaks in fingerprints  $F_{il}$  (Lemma 5.1). For the honesty metric  $\beta_i$  of a legitimate node, this peak location should be normally distributed (subject to noise, measurement error) around the reported location. For a spoofer that reports a random location, the peak location is uniformly distributed. A similar (but inverse) argument holds for  $\gamma_{ij}$ . Hence, we simply need to show is that the definitions of  $\beta_i$  and  $\gamma_i$  which are both products of the form  $g(X)$  can be bounded in expectation if  $X$  is uniform or normally distributed.

To this end, consider two random variables  $u$  and  $v$  which are respectively uniform and normally distributed between 0 and  $2\pi$  with mean 0 and variance  $\sigma^2$ . Let  $S = \sqrt{2}\sigma(\ln \frac{1}{\sigma})^{0.5}$ , the value at which the minimization in  $g(x)$  is triggered.  $E[g(v)]$  and  $E[g(u)]$  are as follows:

$$\begin{aligned} E[g(v)] &= \int_{-S}^S f(x; 0, \sigma^2) dx + \sqrt{8\pi} \int_{-\infty}^{-S} [f(x; 0, \sigma^2)]^2 dx \\ &\geq \int_{-S}^S f(x; 0, \sigma^2) dx = \text{erf} \left( \frac{S}{\sigma\sqrt{2}} \right) \geq 1 - \sigma \end{aligned} \quad (6)$$

Where  $\text{erf}(\cdot)$  is the well known Error function and using  $1 - \text{erf}(x) < e^{-x^2}$ . Similarly, we can evaluate  $E[u(n)]$  as:

$$\begin{aligned} E[g(u)] &= \int_{-S}^S \frac{1}{2\pi} dx + 2\sqrt{2\pi} \int_{-2\pi}^{-S} \frac{1}{2\pi} f(x; 0, \sigma^2) dx \\ &\leq \frac{S}{\pi} + \frac{1}{\sqrt{2\pi}} \left( 1 - \text{erf} \left( \frac{S}{\sigma\sqrt{2}} \right) \right) \leq \sqrt{\sigma}\kappa \end{aligned} \quad (7)$$

By assumptions A.1-A.3, we can apply these bounds to write the expectation of the honesty metric  $\beta_i$  as a product of those of the independent variables:

$$E[\beta_{spoof}] = \prod_l E[g(u; 0, \hat{\sigma}_\phi^2)] E[g(u; 0, \hat{\sigma}_\theta^2)] \leq \left[ \sqrt{\hat{\sigma}_\theta \hat{\sigma}_\phi} \kappa \right]^m$$

$$E[\beta_{legit}] = \prod_l E[g(v; 0, \hat{\sigma}_\phi^2)] E[g(v; 0, \hat{\sigma}_\theta^2)] \geq 1 - m\hat{\sigma}_\theta\hat{\sigma}_\phi$$

Applying a similar argument, the similarity metric  $\gamma$  is:

$$E[\gamma_{spoof}] = \prod_{p=1}^k E[f(v; 0, 2\sigma_\phi^2)] f(v; 0, 2\sigma_\theta^2) \geq 1 - 2mk\sigma_\theta\sigma_\phi$$

$$E[\gamma_{legit}] = \prod_{p=1}^k E[g(u; 0, 2\sigma_\phi^2)] g(u; 0, 2\sigma_\theta^2) \leq \left[ \sqrt{2\sigma_\theta\sigma_\phi} \kappa \right]^{mk}$$

Combining the above equations, we prove Eqn. 5.  $\square$

A natural question one might ask is if the above lemma holds in general environments, where its assumptions A.1-A.3 may be too stringent. Our extensive experimental results in Sec. VII show that our bounds on  $\alpha$  approximately predict performance in general environments. Further, Sec. §VII-A shows that results from an anechoic chamber, which emulate free-space conditions where the lemma's assumptions can be directly enforced, tightly follow the bounds of Lemma 5.1.

In sum, one can adopt the above lemma to distinguish adversarial nodes from legitimate nodes, purely based on  $\alpha$ .

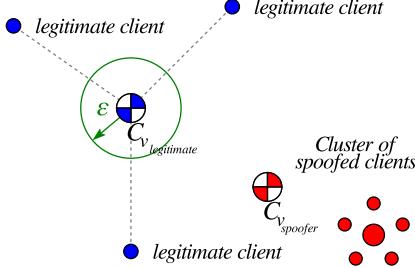


Fig. 6: **Coverage guarantee:** An  $\epsilon$  ball around the ground-truth centroid,  $C_{V_{\text{legitimate}}}$ . Theorem 6.1 finds  $\epsilon(\mathcal{P})$  so that server positions remain within this ball even in the presence of spoofed clients.

However, an interesting alternative is to incorporate  $\alpha$  directly into multi-robot controllers to give provable service guarantees to legitimate nodes. The next section show how  $\alpha_i$  readily integrates with robotic coverage controllers, in particular.

## VI. THREAT-RESISTANT DISTRIBUTED CONTROL

This section describes how our spoof detection method from Sec. §V integrates with well-known coverage controllers from [5, 29, 30]. The area coverage problem deals with the positioning server robots to minimize Euclidean distance to certain areas of interest in the environment. These areas are determined by an importance function  $\rho(q)$  that is defined over the environment  $\mathcal{Q} \subset \mathbb{R}^3$  of size  $L(\mathcal{Q})$ . For our coverage problem, the peaks of the importance are determined by client positions  $P$ , for e.g.,  $\rho(q, P) = \rho_1(q) + \dots + \rho_c(q)$  where  $\rho_i(q)$  quantifies the influence of client  $i$ 's position on the importance function. Using [5, 29, 30], server robot positions optimizing coverage over  $\rho(q, P)$  will minimize their distance to clients.

To account for spoofed clients, we modify the importance function  $\rho(q, P)$  using the  $\alpha_i$  for each client  $i \in [c]$  that is computed by Algorithm 1. For e.g., we can multiply each client-term in  $\rho(q, P)$  by its corresponding confidence weight:  $\rho(q, P)_\alpha = \alpha_1\rho_1(q) + \dots + \alpha_c\rho_c(q)$ . Given the properties of these weights derived in Theorem 5.4, ie.  $\alpha_i$  is bounded near zero for a spoofed client and near one for a legitimate client, the effect of multiplication by the  $\alpha$ 's is that terms corresponding to spoofed clients will be bounded to a small value (see Fig. 6); providing resilience to the spoofing attack.

For simplicity, we assume the importance function  $\rho(q)$  is static (from [5]) and  $\alpha$ 's from Algorithm 1 are computed once, at the beginning of the coverage algorithm. We note that our approach readily extends to the adaptive case in [29, 30] when the importance function (and location of clients) change, by having the service robots exchange their learned importance function. This in turn can trigger a re-calculation of  $\alpha$  values.

We now show that computed server positions are impacted by spoofers to within a closed-form bound, that depends on problem parameters like signal-to-noise ratio. Theorem 6.1 below solves Problem 2 of our problem statement (Sec. §III).

**Theorem 6.1:** Let  $X$  be a set of server robot positions and  $P = S \cup \tilde{S}$  be a set of client positions where  $S$  the set of spoofed client positions, and  $\tilde{S}$  is the set of legitimate clients. The identities of the clients being spoofed is assumed

unknown. Let  $\{\alpha_1, \dots, \alpha_c\}$  be a set of confidence weights satisfying Theorem 5.4 and assume a known importance function  $\rho(q, P) = \rho_1(q) + \dots + \rho_c(q)$  that is defined over the environment  $\mathcal{Q} \subset \mathbb{R}^3$  of size  $L(\mathcal{Q})$ . Define  $C_V = \{x_1^*, \dots, x_m^*\}$  to be the set of server positions optimized over  $\rho(q, \tilde{S})$ , ie. where there are zero spoofed clients and  $C_{V_\alpha}$  to be the set of server positions optimized over  $\rho(q, P)_\alpha = \alpha_1\rho_1(q) + \dots + \alpha_c\rho_c(q)$  where there is at least one spoofed client, ie.  $|S| \geq 1$ . If  $\{\alpha_1, \dots, \alpha_c\}$  satisfy Theorem 5.4, we have that  $\forall x \in C_{V_\alpha}$  there exists a unique  $y \in C_V$ , where in the expected case  $\text{dist}(x, y) \leq \epsilon(m, s, \sigma_\phi, \sigma_\theta, \kappa)$

$$\epsilon = \max \left\{ [\sqrt{\hat{\sigma}_\theta \hat{\sigma}_\phi} \kappa]^m [2mk\sigma_\theta \sigma_\phi]^s, cm\hat{\sigma}_\theta \hat{\sigma}_\phi [\sqrt{2\sigma_\theta \sigma_\phi} \kappa]^{mk} \right\} L(\mathcal{Q})$$

and  $m, s, \sigma_\phi, \sigma_\theta, \kappa$  are problem parameters as in Theorem 5.4.

*Proof:* We make an important observation that  $E[\alpha_i] \leq a$  if client  $i$  is a spoofed node, and  $E[\alpha_i] \geq b$  otherwise; hence:

$$\rho(q, P)_\alpha = a(\rho_1(q) + \dots + \rho_s(q)) + b(\rho_{s+1}(q) + \dots + \rho_c(q))$$

is the maximal effect that the presence of spoofed clients can have on the importance function. Intuitively, all spoofed clients have a weight of *at maximum*  $a$  and all legitimate clients have a reduced weight of *at minimum*  $b$ . Using this observation we can bound the influence of the spoofed clients on computed router control inputs (see Fig. 6). Specifically, recall from [5] that the position control for each server is:  $u_l = -2M_V(C_V - c_l)$ , where  $M_V = \int_V \rho(q)dq$ ,  $C_V = \frac{1}{M_V} \int_V q\rho(q)dq$  and  $V$  is the voronoi partition for router  $l$  defined as all points  $q \in \mathcal{Q}$  with  $\text{dist}(q, x_l) < \text{dist}(q, x_g)$  where  $g \neq l$ . Using the importance function from above we can write  $C_{V_\alpha} = \frac{1}{M_{V_\alpha}}(aC_{V_S} + bC_{V_L})$  where  $C_{V_S}$  is the component of the centroid computed over spoofed nodes and  $C_{V_L}$  is the component of the centroid computed over legitimate nodes and  $M_{V_\alpha}$  is defined shortly. We rewrite  $C_{V_S}$  as a perturbation of the centroid over legitimate nodes as  $C_{V_S} = C_{V_L} + \vec{v}\|\vec{e}\|$  where  $\vec{v}$  is an arbitrary unit vector and the magnitude of  $\vec{e}$  can be as large as the length of the operative environment,  $\|\vec{e}\| \leq L(\mathcal{Q})$ . Let the total mass be  $T = M_{V_S} + M_{V_L}$ . We can write a similar expression for the mass  $M_{V_\alpha}$  using the bounds  $a$  and  $b$  as  $M_{V_\alpha} = bT + (a-b)M_{V_L}$ . Substituting these expressions into  $C_{V_\alpha}$  and simplifying gives  $C_{V_\alpha} = \frac{C_{V_L} + b\vec{v}\|\vec{e}\|}{bT + (a-b)M_{V_L}}$ . Combining this expression with the router control input:

$$u_l = k([a + b]C_{V_L} - p_l) + b\|\vec{e}\|\vec{v} \quad (8)$$

Where  $k = -2(bT + aM_{V_L})$ . If  $(a + b) = 1$ , this control input drives the server robot  $l$  to a neighborhood of size  $\epsilon = b\|\vec{e}\| \leq bL(\mathcal{Q})$  centered around the centroid  $C_L$  defined over the legitimate clients. So if  $b = \max \left\{ [\sqrt{\hat{\sigma}_\theta \hat{\sigma}_\phi} \kappa]^m [2mk\sigma_\theta \sigma_\phi]^s, cm\hat{\sigma}_\theta \hat{\sigma}_\phi [\sqrt{2\sigma_\theta \sigma_\phi} \kappa]^{mk} \right\}$  from Theorem 5.4 Equation (5), then:

$$\epsilon = \max \left\{ [\sqrt{\hat{\sigma}_\theta \hat{\sigma}_\phi} \kappa]^m [2mk\sigma_\theta \sigma_\phi]^s, cm\hat{\sigma}_\theta \hat{\sigma}_\phi [\sqrt{2\sigma_\theta \sigma_\phi} \kappa]^{mk} \right\} L(\mathcal{Q})$$

then we have  $(a + b) = 1$  as desired, proving the lemma.  $\square$

## VII. EXPERIMENTAL RESULTS

This section describes our results from an experimental evaluation of our theoretical claims. Our aerial servers were

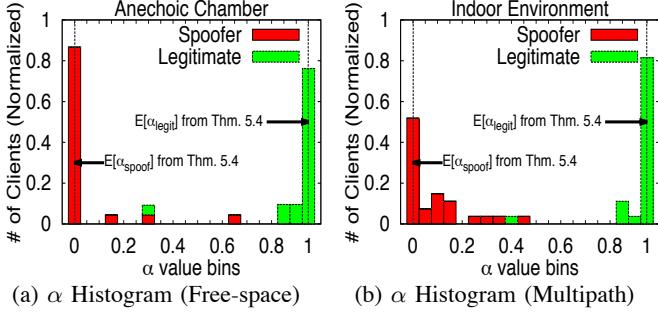
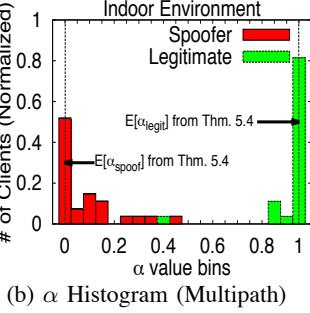
(a)  $\alpha$  Histogram (Free-space)(b)  $\alpha$  Histogram (Multipath)

Fig. 7: **Experimental Evaluation of  $\alpha$ :** (a) In an anechoic chamber approximating our assumptions A.1-A.3 (§5.4),  $\alpha$  agrees with theoretical expectations. (b) in a typical multipath environment, experimental results largely follow theoretical predictions. Data shows that  $\alpha = 0.5$  is a good threshold value.

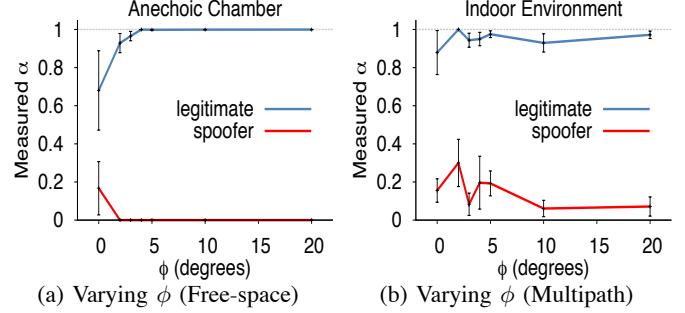
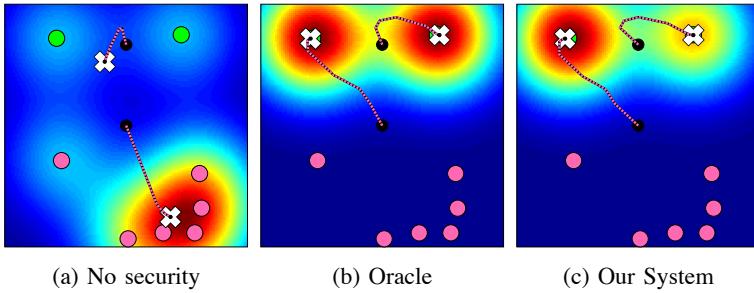
(a) Varying  $\phi$  (Free-space)(b) Varying  $\phi$  (Multipath)

Fig. 8: **Co-Aligned Clients:** We vary the angle  $\phi$  between a legitimate and malicious client, relative to a single server and plot  $\alpha$ , in (a) an anechoic chamber and (b) an indoor environment. The minimum  $\phi$  needed to distinguish the clients is only: (a)  $3^\circ$  in freespace, (b)  $0^\circ$  in multipath settings.



(a) No security

(b) Oracle

(c) Our System

Fig. 9: **Experimental Results for Sybil Attack in Multi-Agent Coverage:** Depicts the total distance of converged quadrotor server positions (white  $\times$ ) to legitimate clients (●) and six spoofed clients (●). We compare this quantity for: (a) A system with no security, where each spoofed clients create a false peak in importance function, (b) ground truth importance function, and (c) our system where applying  $\alpha$  weights from Algorithm 1 recovers the true importance function. (d) depicts the resulting ground-truth cost computed with respect to legitimate clients as spoofed nodes are introduced to the system. The red dotted line shows that our system performs close to ground truth even as spoofed clients comprise more than twice the network.

implemented on two AscTec Atomboard computing platforms, equipped with Intel 5300 Wi-Fi cards with two antennas each, mounted on two AscTec Hummingbird quadrotors. Our clients were ten iRobot Create robots, each equipped with Asus EEEPC netbooks and single-antenna Wi-Fi cards. An adversarial client forged multiple identities by spawning multiple packets containing different identities (up to 75% of the total number of legitimate clients in the system), and could use a different transmit power for each identity. The adversary advertised identities by modifying the Wi-Fi MAC field, a common technique for faking multiple identities [31].

**Evaluation:** We evaluate our system in two environments: (1) An indoor environment equipped with a Vicon motion capture system to aid quadrotor navigation; (2) An anechoic chamber to emulate a free-space setting that closely models assumptions A.1-A.3 in Sec. §V. We estimated theoretical expected standard deviations  $\sigma_\theta, \sigma_\phi$  of about 0.7 degrees (see Lemma 5.1). After including the standard deviation in reported location, based on the known errors of our localization framework, this

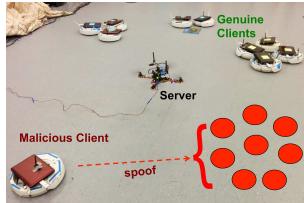


Fig. 10: Experimental setup

increased to  $\hat{\sigma}_\theta, \hat{\sigma}_\phi$  of about 2 degrees (see Theorem 5.4). We compare our system against a baseline that identifies fake clients by comparing their Received Signal Strength Indicators (RSSI) against other clients in the network, akin to [27].

**Roadmap:** We conduct three classes of experiments: (1) Microbenchmarks to validate our client confidence metric, both in free-space and multipath indoor environments (Sec. §VII-A). (2) Experiments applying this confidence metric to quarantine adversaries (Sec. §VII-B). (3) Application of our system to secure the coverage problem against Sybil attacks (Sec. §VII-C).

#### A. Microbenchmarks on the Confidence Metric

This experiment studies the correctness of our system's confidence metric  $\alpha$ . Recall from theory in §V that  $\alpha$ 's measured by a server robot distinguish between unique clients based on their diverse physical directions and the presence of multipath reflections. Thus, a free-space environment (i.e. with no multipath) is particularly challenging to our system.

**Method:** To approximate free-space, we measured  $\alpha$ 's values in a radio-frequency anechoic chamber which attenuates reflected paths by about 60 dB, for a legitimate and malicious client from one server robot 12 m away. Next, in a 10 m  $\times$  8 m indoor room (a typical multipath case), we measured  $\alpha$ 's from one

server for up to ten legitimate clients and ten spoofed clients.

**Results:** In Fig. 7, the values of  $\alpha$  in the anechoic chamber tightly follow our theoretical bounds in Theorem 5.4 (Fig. 8(c)). As expected, our results in indoor multipath environments exhibit a larger variance but follow the trend suggested by theory. Further, we stress our confidence metric by isolating the case of colinearity in both environments. In Fig. 8, we consider a spoofing adversary initially co-aligned with a legitimate client, and measure  $\alpha$  as the angle of separation,  $\phi$ , is increased from  $0^\circ$  to  $20^\circ$  relative to the server robot. In the anechoic chamber at  $\phi$  close to  $0^\circ$ , the fingerprints of both the legitimate and adversarial nodes are virtually identical, each with precisely one peak at  $0^\circ$ . Consequently,  $\alpha$  for the legitimate node is much below 1, indicating that it is believed to be adversarial (i.e. the term  $1 - \gamma$  in  $\alpha$  approaches 0 in Eqn. 2). However,  $\alpha$  for the legitimate client quickly approaches 1, even if  $\phi = 3^\circ$  in the anechoic chamber. In fact,  $\alpha$  is virtually identical to 1 beyond  $10^\circ$ , indicating that a single server robot can distinguish closely aligned legitimate and adversarial clients even in free-space. Fig. 8b shows that multipath can distinguish clients even at  $\phi = 0^\circ$ , due to additional reflected paths that help disambiguate these clients.

### B. Performance of Sybil Attack Detection

In this experiment, we measure our system’s classification performance on legitimate and spoofed clients, in the presence of static, mobile, and power-scaling adversaries.

**Method:** Each run consisted of one quadrotor server, and (randomly positioned) ten control clients, or nine legitimate clients with an adversary reporting two to nine spoofed clients. Each Sybil attack was performed under three modalities: (1) a stationary attacker with a fixed transmission power, (2) a mobile attacker (random-walk and linear movements), and (3) an attacker scaling the per-packet power by a different amount for each spoofed client, from 1 to 31 mW. The quadrotor server classifies clients with an  $\alpha < 0.5$  as spoofed (see Fig. 7). The baseline Received Signal Strength Indicator (RSSI) classifier uses a 2 dB thresholded minimum dissimilarity, a technique previously applied in static networks [27, 35].

	Our System		RSSI	
	TPR	FPR	TPR	FPR
Stationary	96.3	3.0	81.5	9.1
Mobile	96.3	6.1	85.2	6.1
Power-scaling	100.0	3.0	74.1	27.3

**TABLE I: Classification performance:** True positive rates (TPR) and false positive rates (FPR) when our system classifies clients with  $\alpha < 0.5$  as spoofed, compared against a Received Signal Strength (RSSI) baseline. We perform experiments across many robot client-server topologies for 3 classes of adversaries — stationary, mobile, and adversaries scaling power differently for each spoofed client.

**Results:** Table I summarizes our results for each modality. Compared to the RSSI classifier, our technique exhibits a high true positive and low false positive rate of about 96% and 4%, across multiple network topologies. In particular, because our classifier computes  $\alpha$  using the ratio of wireless signal channels (Sec. §IV), it is robust to power-scaling Sybil attacks

where RSSI performs poorly. Our solution exhibits consistent performance in both power-scaling and mobile scenarios.

### C. Application to Multi-Agent Coverage

We implement the multi-agent coverage problem from [5], where a team of aerial servers position themselves to minimize distance to client robots at reported positions  $p_i, i \in [c]$ . We use an importance function  $\rho(q, P) = \rho_1(q) + \dots + \rho_c(q)$  defined in Section VI where each client term is a Gaussian-shaped function  $\rho_i(q) = \exp(-\frac{1}{2}(q - p_i)^T(q - p_i))$  (Fig. 9b). An  $\alpha$ -modified importance function is implemented as  $\rho(q, P)_\alpha = \alpha_1\rho_1(q) + \dots + \alpha_c\rho_c(q)$  where the  $\alpha$  terms are computed using Algorithm 1 (Fig. 9c).

**Method.** For each experiment we randomly place three clients  $i, j$ , and  $k$  in an 8 m x 10 m room along with two AscTec quadrotor servers. Fig. 9(a)-(c) show an example client-router topology, with an adversary spoofing six Sybil clients. We measure the total distance of the routers upon convergence from their optimal converged locations in three scenarios: (1) a naive system with no cyber-security; (2) our system; (3) an oracle that discards Sybil clients a priori.

**Results:** Fig. 9(a)-(c) depicts the converged locations for a system with no security, an oracle, and our system in a candidate topology. We observe that our system approximates oracle performance, by incorporating  $\alpha$  weights in our controller. Fig. 9d demonstrates the ability of our system to bound the cost near optimal even as spoofers enter the network (comprising up to 300%).

**Aggregate Results:** Across multiple topologies and 12 separate runs, the maximum distance from each quadrotor to the oracle solution is on average 3.77 m (stdev: 0.86), in contrast our system achieves proximity to oracle positions of 0.02 m (stdev: 0.02).

## VIII. DISCUSSION

We make the following observations and suggestions for future work: (1) We note that many of the concepts described in this paper are applicable to servers as well, since they also communicate wirelessly. We leave this an interesting problem for future work. (2) Our current implementation runs SAR by making the quadrotor perform a single spin in place. However, we believe it will be interesting for future implementations to perform SAR using other forms of movement, say, along linear paths.

## IX. CONCLUSION

In this paper, we develop a new system to guard against the Sybil attack in multi-robot networks. We derive theoretical guarantees on the performance of our system, that are validated experimentally. While this paper has focused on coverage, it can be readily extended to guard against the Sybil attack in other multi-robot contexts, e.g. unmanned delivery [19], search-and-rescue [20] and formation control [36]. Beyond the Sybil attack, this paper reveals the promise of using the physics of wireless signals as the basis for holistic cyber-security in multi-robot networks against a wide-range of attacks.

## REFERENCES

- [1] Amazon prime air. URL <http://www.amazon.com/b?node=8037720011>.
- [2] Fadel Adib, Swarun Kumar, Omid Aryan, Shyamnath Gollakota, and Dina Katabi. Interference Alignment by Motion. MOBICOM, 2013.
- [3] R.W. Beard, T.W. McLain, D.B. Nelson, D. Kingston, and D. Johanson. Decentralized cooperative aerial surveillance using fixed-wing miniature uavs. *Proceedings of the IEEE*, 94(7):1306–1324, July 2006. ISSN 0018-9219. doi: 10.1109/JPROC.2006.876930.
- [4] Airlie Chapman, Marzieh Nabi-Abdolyousefi, and Mehran Mesbahi. Identification and infiltration in consensus-type networks. *1st IFAC Workshop on Estimation and Control of Networked Systems*, 2009.
- [5] J. Cortes, S. Martinez, T. Karatas, and F. Bullo. Coverage control for mobile sensing networks. 20(2), 2004.
- [6] K. Daniel, B. Dusza, A. Lewandowski, and C. Wietfeld. Airshield: A system-of-systems muav remote sensing architecture for disaster response. In *Systems Conference, 2009 3rd Annual IEEE*, pages 196–200, March 2009. doi: 10.1109/SYSTEMS.2009.4815797.
- [7] JohnR. Douceur. The sybil attack. In Peter Druschel, Frans Kaashoek, and Antony Rowstron, editors, *Peer-to-Peer Systems*, volume 2429 of *Lecture Notes in Computer Science*, pages 251–260. Springer Berlin Heidelberg, 2002. ISBN 978-3-540-44179-3. doi: 10.1007/3-540-45748-8\_24. URL [http://dx.doi.org/10.1007/3-540-45748-8\\_24](http://dx.doi.org/10.1007/3-540-45748-8_24).
- [8] Patrick J. Fitch. *Synthetic Aperture Radar*. Springer, 1988.
- [9] H. Gazzah and S. Marcos. Directive antenna arrays for 3d source localization. In *Signal Processing Advances in Wireless Communications, 2003. SPAWC 2003. 4th IEEE Workshop on*, pages 619–623, June 2003. doi: 10.1109/SPAWC.2003.1319035.
- [10] Houcem Gazzah and Sylvie Marcos. Cramer-Rao bounds for antenna array design. *IEEE Transactions on Signal Processing*, 54:336–345, 2006. doi: 10.1109/TSP.2005.861091.
- [11] Stephanie Gil, Swarun Kumar, Dina Katabi, and Daniela Rus. Adaptive Communication in Multi-Robot Systems Using Directionality of Signal Strength. ISRR, 2013.
- [12] Andrea Goldsmith. *Wireless Communications*. Cambridge University Press, 2005.
- [13] Monson H. Hayes. *Statistical Digital Signal Processing and Modeling*. John Wiley & Sons, Inc., New York, NY, USA, 1st edition, 1996. ISBN 0471594318.
- [14] Fiona Higgins, Allan Tomlinson, and Keith M. Martin. Threats to the swarm: Security considerations for swarm robotics. *International Journal on Advances in Security*, 2, 2009.
- [15] Dongxu Jin and JooSeok Song. A traffic flow theory aided physical measurement-based sybil nodes detection mechanism in vehicular ad-hoc networks. In *Computer and Information Science (ICIS), 2014 IEEE/ACIS 13th International Conference on*, pages 281–286, June 2014. doi: 10.1109/ICIS.2014.6912147. URL [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=6912147&tag=1](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6912147&tag=1).
- [16] Helmut Klausing. Feasibility of a sar with rotating antennas (rosar). In *Microwave Conference, 1989*, 1989.
- [17] Swarun Kumar, Stephanie Gil, Dina Katabi, and Daniela Rus. Accurate indoor localization with zero start-up cost. In *Proceedings of the 20th Annual International Conference on Mobile Computing and Networking*, MobiCom ’14, pages 483–494, New York, NY, USA, 2014. ACM. ISBN 978-1-4503-2783-1. doi: 10.1145/2639108.2639142. URL <http://doi.acm.org/10.1145/2639108.2639142>.
- [18] Swarun Kumar, Ezzeldin Hamed, Dina Katabi, and Li Erran Li. Lte radio analytics made easy and accessible. In *Proceedings of the 2014 ACM Conference on SIGCOMM*, SIGCOMM ’14, pages 211–222, New York, NY, USA, 2014. ACM. ISBN 978-1-4503-2836-4. doi: 10.1145/2619239.2626320. URL <http://doi.acm.org/10.1145/2619239.2626320>.
- [19] Aleksandr Kushleyev, Brian MacAllister, and M. Likhachev. Planning for landing site selection in the aerial supply delivery. In *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*, pages 1146–1153, Sept 2011. doi: 10.1109/IROS.2011.6094840.
- [20] Lanny Lin and Michael A Goodrich. Uav intelligent path planning for wilderness search and rescue. In *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*, pages 709–714. IEEE, 2009.
- [21] Hongbo Liu, Yan Wang, Jian Liu, Jie Yang, and Yingying Chen. Practical user authentication leveraging channel state information (csi). In *Proceedings of the 9th ACM Symposium on Information, Computer and Communications Security*, ASIA CCS ’14, pages 389–400, New York, NY, USA, 2014. ACM. ISBN 978-1-4503-2800-5. doi: 10.1145/2590296.2590321. URL <http://doi.acm.org/10.1145/2590296.2590321>.
- [22] M. MalmirChegini and Y. Mostofi. On the spatial predictability of communication channels. *Wireless Communications, IEEE Trans.*, 11(3), 2012.
- [23] Cherian P. Mathews and Michael D. Zoltowsk. Signal subspace techniques for source localization with circular sensor arrays. Purdue University TechReport, 1994.
- [24] J. Newsome, E. Shi, D. Song, and A. Perrig. The sybil attack in sensor networks: analysis defenses. In *Information Processing in Sensor Networks, 2004. IPSN 2004. Third International Symposium on*, pages 259–268, April 2004. doi: 10.1109/IPSN.2004.1307346.
- [25] R. Olfati-Saber and R.M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *Automatic Control, IEEE Transactions on*, 49(9): 1520–1533, Sept 2004. ISSN 0018-9286. doi: 10.1109/TAC.2004.834113.
- [26] Lynne E. Parker. Distributed algorithms for multi-robot

observation of multiple moving targets. *Autonomous Robots*, 12, 2002.

[27] Jr. Pires, W.R., T.H. de Paula Figueiredo, H.C. Wong, and A.A.F. Loureiro. Malicious node detection in wireless sensor networks. In *Parallel and Distributed Processing Symposium, 2004. Proceedings. 18th International*, pages 24–, April 2004. doi: 10.1109/IPDPS.2004.1302934.

[28] I. Sargeant and A. Tomlinson. Modelling malicious entities in a robotic swarm. In *Digital Avionics Systems Conference (DASC), 2013 IEEE/AIAA 32nd*, Oct 2013.

[29] M. Schwager, Brian J. Julian, and D. Rus. Optimal coverage for multiple hovering robots with downward facing cameras. In *Robotics and Automation, 2009. ICRA '09. IEEE International Conference on*, pages 3515–3522, May 2009. doi: 10.1109/ROBOT.2009.5152815.

[30] Mac Schwager, Daniela Rus, and Jean-Jacques Slotine. Decentralized, adaptive coverage control for networked robots. *The International Journal of Robotics Research*, 28(3):357–375, 2009. URL <http://ijr.sagepub.com/content/28/3/357.abstract>.

[31] Yong Sheng, K. Tan, Guanling Chen, D. Kotz, and A. Campbell. Detecting 802.11 mac layer spoofing using received signal strength. In *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, pages –, April 2008. doi: 10.1109/INFOCOM.2008.239. URL [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=4509834&tag=1](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4509834&tag=1).

[32] Petre Stoica and Nehorai Arye. Music, maximum likelihood, and cramer-rao bound. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 37(5):720–741, May 1989. ISSN 0096-3518. doi: 10.1109/29.17564.

[33] D. Tse and P. Vishwanath. *Fundamentals of Wireless Communications*. Cambridge University Press, 2005.

[34] Jue Wang and Dina Katabi. Dude, where's my card?: Rfid positioning that works with multipath and non-line of sight. *SIGCOMM*, 2013.

[35] Ting Wang and Yaling Yang. Analysis on perfect location spoofing attacks using beamforming. In *INFOCOM, 2013 Proceedings IEEE*, pages 2778–2786, April 2013. doi: 10.1109/INFCOM.2013.6567087. URL [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=6567087](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6567087).

[36] Xiaohua Wang, Vivek Yadav, and SN Balakrishnan. Cooperative uav formation flying with obstacle/collision avoidance. *Control Systems Technology, IEEE Transactions on*, 15(4):672–679, 2007.

[37] Yong Wang, G. Attebury, and B. Ramamurthy. A survey of security issues in wireless sensor networks. *Communications Surveys Tutorials, IEEE*, 8(2):2–23, Second 2006. ISSN 1553-877X. doi: 10.1109/COMST.2006.315852.

[38] Jie Xiong and Kyle Jamieson. Securearray: Improving wifi security with fine-grained physical-layer information. In *Proceedings of the 19th Annual International Conference on Mobile Computing & Networking, MobiCom '13*, pages 441–452, New York, NY, USA, 2013. ACM. ISBN 978-1-4503-1999-7. doi: 10.1145/2500423.2500444. URL <http://doi.acm.org/10.1145/2500423.2500444>.

1145/2500423.2500444.

[39] Jie Yang, Yingying Chen, W. Trappe, and J. Cheng. Detection and localization of multiple spoofing attackers in wireless networks. *Parallel and Distributed Systems, IEEE Transactions on*, 24(1):44–58, Jan 2013. ISSN 1045-9219. doi: 10.1109/TPDS.2012.104.

[40] Zhimin Yang, E. Ekici, and Dong Xuan. A localization-based anti-sensor network system. In *INFOCOM 2007. 26th IEEE International Conference on Computer Communications. IEEE*, pages 2396–2400, May 2007. doi: 10.1109/INFCOM.2007.288. URL [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=4215870](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4215870).